



TITLE:

# クリーク分割問題に対する疎な定式化 (最適化手法の理論と応用の繋がり)

AUTHOR(S):

宮内, 敦史; 鮎川, 矩義

---

CITATION:

宮内, 敦史 ...[et al]. クリーク分割問題に対する疎な定式化 (最適化手法の理論と応用の繋がり). 数理解析研究所講究録 2013, 1829: 149-155

ISSUE DATE:

2013-03

URL:

<http://hdl.handle.net/2433/194801>

RIGHT:

## クリーク分割問題に対する疎な定式化

東京工業大学 大学院社会理工学研究科

宮内 敦史 (Miyuchi Atsushi)

鮎川 矩義 (Sukegawa Noriyoshi)

Graduate School of Decision Science and Technology,  
Tokyo Institute of Technology

### 概要

クリーク分割問題は、クラスタリングをモデル化する際に用いられるグラフ上の最適化問題である。質的データ解析、マイクロアレイ解析、コミュニティ検出、グループテクノロジーなど、実社会への数多くの応用を持つ。しかしながら、クリーク分割問題は NP-困難であり、Grötschel と Wakabayashi により提案された 0-1 整数計画問題としての標準的な定式化は、膨大な不等式制約を抱えている。本稿では、クリーク分割問題に対する彼らの定式化を見直し、新しい定式化を与える。具体的には、簡単な検査で見つけられる冗長な不等式制約のクラスを示し、これをあらかじめ取り除くことを提案する。計算実験では、計算資源の節約だけでなく、計算時間の短縮も確認された。

## 1 はじめに

クリーク分割問題は、クラスタリングをモデル化する際に用いられるグラフ上の最適化問題である。質的データ解析 [3, 6], マイクロアレイ解析 [9], コミュニティ検出 [2, 1], グループテクノロジー [8, 10] など、実社会への数多くの応用を持つ。Zahn の問題（二項関係を同値関係に近似する問題）や Règnier の問題（複数の同値関係を一つに集約する問題）といった、NP-困難性が示された古典的な最適化問題の拡張としても知られている。

クリーク分割問題に対する標準的な定式化は、Grötschel と Wakabayashi [6] により提案された 0-1 整数計画問題としてのものである。厳密解法や発見的解法を設計する際にしばしば用いられているが、膨大な不等式制約を抱えているという問題がある。例えば、頂点数が 100 のとき、制約式は約 50 万本にもなる。したがって、入力が少し大きなグラフになると、計算資源の確保が困難となる。

本稿では、彼らの定式化を見直し、制約式の本数を大幅に減らした定式化を与える。具体的には、簡単な検査で見つけられる冗長な不等式制約のクラスを示し、これをあらかじめ取り除くことを提案する。このクラスに属す制約式は、元問題の線形緩和問題においても冗長であることが示される。計算実験では、計算資源の節約だけでなく、計算時間の短縮も確認された。

本稿で示す結果は、モジュラリティ最大化問題に対する Dinh と Thai [4] の結果の一般化となっている。証明の方針は、本質的には彼らのものと同じである。しかしながら、適用範囲が大幅に広がった点と、彼らが示した冗長な不等式制約のクラスを狭義に包含する点に意義がある。本稿の題目にある「疎な定式化」は、彼らが論文中で用いた “sparse formulation” に倣ったものである。

本稿の構成を以下に示す。次の節では、クリーク分割問題の定義を述べ、上記の標準的な定式化について説明する。第 3 節では、本研究の主結果とその証明を与える。第 4 節では、Dinh と Thai [4] が示した結果と本研究の関係について説明する。第 5 節では、簡単な計算実験を行い、本稿で提案する定式化の有効性を確認する。最後に、第 6 節で結論を述べる。

## 2 準備

完全無向グラフ  $G = (V, E)$  において、頂点集合  $V$  のある分割  $\{V_1, V_2, \dots, V_p\}$  が存在して、枝集合  $A \subseteq E$  が

$$A = \bigcup_{l=1}^p \{\{i, j\} \in E \mid i, j \in V_l\}$$

を満たすとき、 $A$  は  $G$  のクリーク分割と呼ばれる [6]。クリーク分割問題とは、正負を問わない枝重み  $c: E \rightarrow \mathbb{R}$  が与えられたとき、重みの総和が最大となるクリーク分割を求める問題である。

ここから、Grötschel と Wakabayashi [6] により提案された定式化について説明する。はじめに、頂点集合を  $V = \{1, 2, \dots, n\}$  とする。次に、枝集合  $A \subseteq E$  に対し、0-1 変数ベクトル  $\mathbf{x} = (x_{ij})_{1 \leq i < j \leq n}$  を対応させる。各成分  $x_{ij}$  は、 $\{i, j\} \in A$  ならば 1 をとり、 $\{i, j\} \notin A$  ならば 0 をとる。「枝集合  $A$  がクリーク分割である」ことと、「任意の  $i, j, k \in V$  について、 $\{i, j\} \in A$  かつ  $\{j, k\} \in A$  ならば  $\{i, k\} \in A$  である」ことの同値性に着目すると、クリーク分割問題は

$$\begin{aligned} (\text{IP}): \quad & \max. \quad \sum_{1 \leq i < j \leq n} c_{ij} x_{ij} \\ \text{s. t.} \quad & x_{ij} + x_{jk} - x_{ik} \leq 1 \quad \forall 1 \leq i < j < k \leq n, \\ & x_{ij} - x_{jk} + x_{ik} \leq 1 \quad \forall 1 \leq i < j < k \leq n, \\ & -x_{ij} + x_{jk} + x_{ik} \leq 1 \quad \forall 1 \leq i < j < k \leq n, \\ & x_{ij} \in \{0, 1\} \quad \forall 1 \leq i < j \leq n, \end{aligned}$$

と定式化できる。ただし、 $c_{ij}$  は枝  $\{i, j\}$  の重みである。また、上の 3 種類の不等式制約は推移性制約と呼ばれる。この定式化では、変数は  $n(n-1)/2$  個、推移性制約は  $n(n-1)(n-2)/2$  本ある。(IP) の 0-1 制約  $x_{ij} \in \{0, 1\}$  を  $x_{ij} \in [0, 1]$  に置き換えて得られる線形緩和問題を (LP) と呼ぶ。

## 3 主結果

本節では、本研究の主結果について述べる。まず、与えられた枝重み  $c: E \rightarrow \mathbb{R}$  に対し、重み为非負である枝の集合  $E_+ = \{\{i, j\} \in E \mid c_{ij} \geq 0\}$  を定義する。我々が提案する新たな定式化は、(IP) から一部の推移性制約を取り除いた

$$\begin{aligned} (\text{IP}_{\text{sparse}}): \quad & \max. \quad \sum_{1 \leq i < j \leq n} c_{ij} x_{ij} \\ \text{s. t.} \quad & x_{ij} + x_{jk} - x_{ik} \leq 1 \quad \forall 1 \leq i < j < k \leq n, \quad \{i, j\} \in E_+ \vee \{j, k\} \in E_+, \\ & x_{ij} - x_{jk} + x_{ik} \leq 1 \quad \forall 1 \leq i < j < k \leq n, \quad \{i, j\} \in E_+ \vee \{i, k\} \in E_+, \\ & -x_{ij} + x_{jk} + x_{ik} \leq 1 \quad \forall 1 \leq i < j < k \leq n, \quad \{j, k\} \in E_+ \vee \{i, k\} \in E_+, \\ & x_{ij} \in \{0, 1\} \quad \forall 1 \leq i < j \leq n, \end{aligned}$$

である。このとき、以下が成り立つ。

**定理 1.** (IP) と  $(\text{IP}_{\text{sparse}})$  は同じ最適解集合を持つ。

**証明.**  $(\text{IP}_{\text{sparse}})$  の任意の最適解

$$\mathbf{x}^* = (x_{ij}^*)_{1 \leq i < j \leq n}$$

が (IP) の実行可能解であることを示せば十分である。ここで、

$$E^* = \{\{i, j\} \mid x_{ij}^* = 1\}$$

とし、 $(V, E^*)$  における連結成分の集合を  $\{(V_1, E_1^*), (V_2, E_2^*), \dots, (V_p, E_p^*)\}$  とする。各  $l = 1, 2, \dots, p$  について  $(V_l, E_l^*)$  が完全であることを示せばよい。以下では、一つの連結成分  $(V_l, E_l^*)$  に注目する。まず、次の補題を示す。

**補題 1.**  $(V_l, E_l^* \cap E_+)$  は連結である。

**証明.**  $V_l$  を任意の二つの頂点集合  $S$  と  $T$  に分割したとき、必ず  $S$  と  $T$  の頂点を各端点に持つ  $E_l^* \cap E_+$  の枝が存在することを示せばよい。 $(V_l, E_l^*)$  の定義より、 $S$  と  $T$  の頂点を各端点に持つ  $E_l^*$  の枝が存在する。ここで、これらの枝のうち  $E_+$  に属するものが一つもない、すなわち、これらの枝の重みがすべて負であると仮定する。すると、 $\mathbf{x}^*$  においてこれらの枝に対応するすべての変数の値を 1 から 0 に変更しても、依然として  $(\text{IP}_{\text{sparse}})$  の実行可能解であるが、その目的関数値は  $\mathbf{x}^*$  のものよりも高くなる。これは  $\mathbf{x}^*$  の最適性に矛盾する。  $\square$

よって、任意の 2 頂点  $i, j \in V_l$  の間に、 $E_l^* \cap E_+$  の枝からなる道  $i = u_0, u_1, \dots, u_q = j$  がとれる。いま、 $\{u_0, u_1\} \in E_+$  (かつ  $\{u_1, u_2\} \in E_+$ ) であるから、 $(\text{IP}_{\text{sparse}})$  は推移性制約

$$x_{u_0 u_1} + x_{u_1 u_2} - x_{u_0 u_2} \leq 1$$

を持つ。そして、 $x_{u_0 u_1}^* = x_{u_1 u_2}^* = 1$  であるため、 $x_{u_0 u_2}^* = 1$  が成り立つ。同様に、 $\{u_2, u_3\} \in E_+$  より、 $(\text{IP}_{\text{sparse}})$  は推移性制約

$$x_{u_0 u_2} + x_{u_2 u_3} - x_{u_0 u_3} \leq 1$$

を持つ。ここで、 $\{u_0, u_2\} \in E_+$  は必ずしも成り立つわけではないことに注意する。これに  $x_{u_0 u_2}^* = x_{u_2 u_3}^* = 1$  を代入すると、 $x_{u_0 u_3}^* = 1$  が得られる。この操作を繰り返すと、最終的に  $x_{u_0 u_q}^* = x_{ij}^* = 1$  が成り立つ。よって、 $(V_l, E_l^*)$  は完全である。  $\square$

$(\text{IP}_{\text{sparse}})$  の 0-1 制約  $x_{ij} \in \{0, 1\}$  を  $x_{ij} \in [0, 1]$  に置き換えて得られる線形緩和問題を  $(\text{LP}_{\text{sparse}})$  と呼ぶ。このとき、以下が成り立つ。

**定理 2.**  $(\text{LP})$  と  $(\text{LP}_{\text{sparse}})$  は同じ最適解集合を持つ。

**証明.** 定理 1 の証明と同様に、 $(\text{LP}_{\text{sparse}})$  の任意の最適解

$$\mathbf{x}^* = (x_{ij}^*)_{1 \leq i < j \leq n}$$

が  $(\text{LP})$  の実行可能解であることを示せば十分である。はじめに、 $\mathbf{x}^*$  の残余ネットワークを

$$\mathbf{d}^* = (d_{ij}^*)_{1 \leq i < j \leq n} = \mathbf{1} - \mathbf{x}^*$$

とおく。すると、 $\mathbf{x}^*$  に関する  $(\text{LP})$  の推移性制約は、 $\mathbf{d}^*$  に関する三角不等式に対応することがわかる。例えば、

$$x_{ij}^* + x_{jk}^* - x_{ik}^* \leq 1 \Leftrightarrow d_{ik}^* \leq d_{ij}^* + d_{jk}^*$$

のように対応付けられる。したがって、 $\mathbf{d}^*$  が三角不等式を満たすことを示せばよい。ここで、

$$\bar{E}^* = \{\{i, j\} \mid d_{ij}^* < 1 (\Leftrightarrow x_{ij}^* > 0)\}$$

とし、 $(V, \bar{E}^*)$  における連結成分の集合を  $\{(V_1, \bar{E}_1^*), (V_2, \bar{E}_2^*), \dots, (V_p, \bar{E}_p^*)\}$  とする。すると、各連結成分内の三角不等式だけを確認すれば十分である。なぜならば、複数の連結成分をまたぐ三角不等式では、そこに登場する項のうち少なくとも二つは 1 になるため、必ず満たされるからである。補題 1 と同様に、以下の主張が成り立つ。証明の方針は同じであるため、ここでは割愛する。

**補題 2.**  $(V_l, \bar{E}_l^* \cap E_+)$  は連結である。

よって、任意の2頂点  $i, j \in V_l$  の間に、 $\bar{E}_l^* \cap E_+$  の枝からなる道がとれる。これらの道のうち、 $\mathbf{d}^*$  を重みとしたときの最短経路を  $i = u_0, u_1, \dots, u_q = j$  とし、その距離を  $d'_{ij}$  とする。いま、 $(\text{LP}_{\text{sparse}})$  の推移性制約、すなわち  $\mathbf{d}^*$  に対する三角不等式を

$$\begin{aligned} d'_{ij} &= d_{u_0 u_1}^* + d_{u_1 u_2}^* + \dots + d_{u_{q-1} u_q}^* \\ &\geq d_{u_0 u_2}^* + d_{u_2 u_3}^* + \dots + d_{u_{q-1} u_q}^* \\ &\geq \dots \geq d_{u_0 u_{q-1}}^* + d_{u_{q-1} u_q}^* \geq d_{u_0 u_q}^* = d_{ij}^*, \end{aligned}$$

と繰り返し適用すれば、 $d'_{ij} \geq d_{ij}^*$  が導かれる。ここで、各2頂点  $i, j \in V_l$  に対する  $d'_{ij}$  をまとめたものを  $\mathbf{d}' = (d'_{ij})_{i < j}$  とする。その構成法から、 $\mathbf{d}'$  は三角不等式を満たすが、実は  $i, j \in V_l$  に関する  $(d_{ij}^*)_{i < j}$  と一致する。これが示せれば、 $i, j \in V_l$  に関する  $(d_{ij}^*)_{i < j}$  が三角不等式を満たすことになるから、証明は完了する。よって、最後にこれを示す。まず、 $\{i, j\} \in E_+$  であれば、 $d'_{ij} \leq d_{ij}^*$  より  $d'_{ij} = d_{ij}^*$  がわかる。よって、ある  $i, j \in V_l$  について  $d'_{ij} \neq d_{ij}^*$  である、すなわち  $d'_{ij} > d_{ij}^*$  が成り立つと仮定すると、 $\{i, j\} \notin E_+$  を満たすことになる。すると、 $(\text{LP}_{\text{sparse}})$  の実行可能解

$$\mathbf{x}' = \begin{cases} 1 - d'_{ij} & i, j \in V_l \text{ のとき,} \\ x_{ij}^* & \text{その他} \end{cases}$$

は  $\mathbf{x}^*$  よりも高い目的関数値をとることになる。これは  $\mathbf{x}^*$  の最適性に矛盾する。 □

定理1と定理2は、(IP)と(LP)の両定式化において、推移性制約のクラス

$$\begin{aligned} x_{ij} + x_{jk} - x_{ik} &\leq 1 \quad \forall 1 \leq i < j < k \leq n, \quad \{i, j\} \notin E_+ \wedge \{j, k\} \notin E_+, \\ x_{ij} - x_{jk} + x_{ik} &\leq 1 \quad \forall 1 \leq i < j < k \leq n, \quad \{i, j\} \notin E_+ \wedge \{i, k\} \notin E_+, \\ -x_{ij} + x_{jk} + x_{ik} &\leq 1 \quad \forall 1 \leq i < j < k \leq n, \quad \{j, k\} \notin E_+ \wedge \{i, k\} \notin E_+, \end{aligned}$$

が冗長であることを主張している。無向グラフ  $G_+ = (V, E_+)$  における頂点  $i$  の次数を  $d_i$  とおくと、 $(\text{IP}_{\text{sparse}})$  と  $(\text{LP}_{\text{sparse}})$  に含まれる推移性制約の本数の上界値は

$$\sum_{1 \leq i < j \leq n} (d_i + d_j) = (n-1) \sum_{1 \leq i \leq n} d_i = O(|E_+|n)$$

で与えられる。制約式の本数が実際にどの程度減少するかについては、第5節の計算実験において確認する。

## 4 関連研究

本節では、Dinh と Thai [4] により示された結果を紹介し、前節で述べた我々の結果との関係について説明する。彼らは、グラフ分割問題の一種であるモジュラリティ最大化問題に対して、我々と同様の疎な定式化を提案している。モジュラリティとは、2004年に Newman と Girvan [7] により提案されたグラフ分割の評価関数である。より正確には、無向グラフ  $G = (V, E)^{*1}$  に対して、頂点集合  $V$  の分割  $\mathcal{C}$  を引数とし、 $-1/2$  以上  $1$  以下の実数値を返す関数である。枝数を  $m$ 、隣接行列の  $(i, j)$  成分を  $A_{ij}$ 、頂点  $i$  の次数を  $d_i$ 、頂点  $i$  が属す分割を  $C_i$  とおくと、モジュラリティは

$$Q(\mathcal{C}) = \frac{1}{2m} \sum_{i, j \in V} \left( A_{ij} - \frac{d_i d_j}{2m} \right) \delta(C_i, C_j)$$

と表される。ここで、 $\delta$  はクロネッカーの記号である。モジュラリティが大きい分割ほど、まとまりらしい分割となることが経験的に知られており、モジュラリティ最大化問題はコミュニティ検出という分野で盛んに研究されている [5]。

<sup>\*1</sup>  $G$  は完全である必要はなく、一般の無向グラフである。

Aloise ら [2] が述べている通り、モジュラリティ最大化問題はクリーク分割問題として定式化できる。第 3 節と同様に、頂点集合を  $V = \{1, 2, \dots, n\}$  とする。そして、頂点  $i$  と  $j$  が同一のコミュニティに属するとき 1, そうでないとき 0 をとる変数を  $x_{ij}$  とおく。すると、最大化する目的関数は

$$\frac{1}{2m} \sum_{1 \leq i, j \leq n} \left( A_{ij} - \frac{d_i d_j}{2m} \right) x_{ij}$$

とおける。任意の頂点  $i$  に対して  $x_{ii} = 1$  であるから、 $x_{ii}$  は変数として採用せず、

$$C = \frac{1}{2m} \sum_{1 \leq i \leq n} \left( A_{ii} - \frac{d_i d_i}{2m} \right)$$

を目的関数に加える。さらに、任意の頂点对  $i, j$  に対して、 $x_{ij} = x_{ji}$  と  $q_{ij} = q_{ji}$  が成り立つから、 $x_{ij}$  ( $i > j$ ) は変数として採用せず、目的関数に 2 を乗じる。以上から、変数として  $x_{ij}$  ( $i < j$ ) のみを採用すると、モジュラリティ最大化問題は

$$\begin{aligned} \max. \quad & \frac{1}{m} \sum_{1 \leq i < j \leq n} \left( A_{ij} - \frac{d_i d_j}{2m} \right) x_{ij} + C \\ \text{s. t.} \quad & x_{ij} + x_{jk} - x_{ik} \leq 1 \quad \forall 1 \leq i < j < k \leq n, \\ & x_{ij} - x_{jk} + x_{ik} \leq 1 \quad \forall 1 \leq i < j < k \leq n, \\ & -x_{ij} + x_{jk} + x_{ik} \leq 1 \quad \forall 1 \leq i < j < k \leq n, \\ & x_{ij} \in \{0, 1\} \quad \forall 1 \leq i < j \leq n, \end{aligned}$$

と定式化できる。この定式化を見ると、クリーク分割問題において

$$c_{ij} = A_{ij} - \frac{d_i d_j}{2m}$$

とした特殊ケースが、モジュラリティ最大化問題であることがわかる。Dinh と Thai [4] は、上の標準的な定式化において、推移性制約のクラス

$$\begin{aligned} x_{ij} + x_{jk} - x_{ik} &\leq 1 \quad \forall 1 \leq i < j < k \leq n, \quad \{i, j\} \notin E \wedge \{j, k\} \notin E, \\ x_{ij} - x_{jk} + x_{ik} &\leq 1 \quad \forall 1 \leq i < j < k \leq n, \quad \{i, j\} \notin E \wedge \{i, k\} \notin E, \\ -x_{ij} + x_{jk} + x_{ik} &\leq 1 \quad \forall 1 \leq i < j < k \leq n, \quad \{j, k\} \notin E \wedge \{i, k\} \notin E, \end{aligned}$$

が冗長であることを示した。

以下では、一番上の制約式を例に、我々の結果との関係について説明する。 $\{i, j\} \notin E$  かつ  $\{j, k\} \notin E$  のとき  $A_{ij} = A_{jk} = 0$  であるため

$$c_{ij} = A_{ij} - \frac{d_i d_j}{2m} < 0 \wedge c_{jk} = A_{jk} - \frac{d_j d_k}{2m} < 0$$

が成り立つ。これは  $\{i, j\} \notin E_+$  かつ  $\{j, k\} \notin E_+$  が成り立つことと同値である。したがって、彼らを与えた冗長な推移性制約のクラスは、我々を与えたものに含まれている。さらに、場合によってはこの包含関係が狭義に成り立つことがわかる。なぜならば、 $\{i, j\} \in E$  であっても、 $\{i, j\} \notin E_+$  となり得るからである。実際、 $A_{ij} = 1$  が満たされても、 $d_i d_j > 2m$  であれば  $c_{ij} < 0$  となる。

## 5 計算実験

本実験では、(IP<sub>sparse</sub>) によって実際にどの程度の推移性制約が取り除かれるのかを確認する。そこで、クリーク分割問題の重要な応用例である同値関係集約問題のインスタンスに対して、(IP) と (IP<sub>sparse</sub>) を適用す

表 1 計算実験の結果

Instance	$n$	(IP)		(IP <sub>sparse</sub> )	
		#constr.	time(sec.)	#constr.	time(sec.)
Wildcats [6]	30	12,180	0.34	10,043	0.31
Cars [6]	33	16,368	0.52	14,708	0.47
Workers [6]	34	17,952	0.66	14,896	0.58
Cetacea [6]	36	21,420	0.61	9,598	0.31
Micro [6]	40	29,640	0.94	22,201	0.73
Soybean [3]	47	48,645	1.58	31,171	1.00
UNO [6]	54	74,412	2.50	45,756	1.59
UN01a [6]	158	1,934,868	114.07	1,161,623	62.87
UN02a [6]	158	1,934,868	114.27	1,542,583	90.24

る。計算実験に用いたコンピュータのスペックは、OS : Windows 7, プロセッサ : Intel Core i5 3.60 GHz, メモリ : 4 GB である。また、実験に使用した最適化ソルバは、Gurobi Optimizer 4.5.0 である。

実験結果を表 1 に示す。第 1 列はインスタンスの名前であり、第 2 列は入力グラフの頂点数である。また、(IP) と (IP<sub>sparse</sub>) の列は、それぞれの定式化の推移性制約の本数と、最適解を得るのに要した計算時間を示している。実際に、(IP) に含まれる多くの推移性制約が (IP<sub>sparse</sub>) では取り除かれていることがわかる。特に、Catacea に対しては 55.2%, UN01a に対しては 40.0% の推移性制約が除去された。さらに、すべてのインスタンスに対して、(IP<sub>sparse</sub>) では (IP) よりも短い時間で最適解が得られている。簡単な実験ではあるが、(IP<sub>sparse</sub>) は省メモリ性と高速性の両面で優れていると言える。

## 6 おわりに

本稿では、クリーク分割問題に対する標準的な定式化を見直し、新たな定式化を提案した。計算実験では、冗長な推移性制約の除去による計算資源の節約が確認された。また、それに伴う計算時間の短縮も見られ、提案した定式化の有効性が伺えた。

今後の研究課題として、以下の二つを考えている。一つは、クリーク分割問題に対する今回の結果の強化である。すなわち、簡単な検査で判定できる冗長な不等式制約のクラスで、より大きなものを与えることである。そしてもう一つは、他の最適化問題に対して、今回のような疎な定式化を提案することである。グラフ上の最適化問題で、その定式化に推移性制約を持つものは多く挙げられるが、今回のような手法が同様に適用できるかどうかは検討する必要がある。

## 参考文献

- [1] G. Agarwal and D. Kempe, Modularity-maximizing graph communities via mathematical programming, *The European Physical Journal B*. 66 (2008) 409–418.
- [2] D. Aloise, S. Cafieri, G. Caporossi, P. Hansen, L. Liberti and S. Perron, Column generation algorithms for exact modularity maximization in networks, *Physical Review E*. 82 (2010) 046112.
- [3] M. J. Brusco and H. F. Köhn, Clustering qualitative data based on binary equivalence relations: neighborhood search heuristics for the clique partitioning problem, *PSYCHOMETRIKA*. 74 (2009)

685–703.

- [4] T. N. Dinh and M. T. Thai, Finding community structure with performance guarantees in complex networks, e-print [arXiv:1108.4034](#) (2011).
- [5] S. Fortunato, Community detection in graphs, *Physics Reports*. 486 (2010) 75–174.
- [6] M. Grötschel and Y. Wakabayashi, A cutting plane algorithm for a clustering problem, *Mathematical Programming*. 45 (1989) 59–96.
- [7] M. E. J. Newman and M. Girvan, Finding and evaluating community structure in networks, *Physical Review E*. 69 (2004) 026113.
- [8] M. Oosten, J. H. G. C. Rutten and F. C. R. Spijksma, The clique partitioning problem: Facets and patching facets, *Networks*. 38 (2001) 209–226.
- [9] H. Wang, B. Alidaee, F. Glover and G. Kochenberger, Clustering of microarray data via clique partitioning, *Journal of Combinatorial Optimization*. 10 (2005) 77–92.
- [10] H. Wang, B. Alidaee, F. Glover and G. Kochenberger, Solving group technology problems via clique partitioning, *International Journal of Flexible Manufacturing Systems*. 18 (2006) 77–97.